

OIPE

RAW SEQUENCE LISTING

DATE: 08/17/2001

PATENT APPLICATION: US/09/760,364

TIME: 12:09:45

Input Set : A:\-41-1.app

Output Set : N:\CRF3\08162001\I760364.raw

3 <110> APPLICANT: Lehmann, Juergen Michael
 4 Shiau, Andrew Kwan-Nan
 5 Tularik Inc.
 7 <120> TITLE OF INVENTION: CAR Modulators: Screening and Treatment of
 8 Hypercholesterolemia
 10 <130> FILE REFERENCE: 018781-004110US
 12 <140> CURRENT APPLICATION NUMBER: US 09/760,364
 13 <141> CURRENT FILING DATE: 2001-01-12
 15 <150> PRIOR APPLICATION NUMBER: US 60/176,398
 16 <151> PRIOR FILING DATE: 2000-01-13
 18 <160> NUMBER OF SEQ ID NOS: 14
 20 <170> SOFTWARE: PatentIn Ver. 2.1
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 348
 24 <212> TYPE: PRT
 25 <213> ORGANISM: Homo sapiens
 27 <220> FEATURE:
 28 <223> OTHER INFORMATION: human constitutive androstane receptor (CAR) alpha
 29 (hCARa)
 31 <400> SEQUENCE: 1
 32 Met Ala Ser Arg Glu Asp Glu Leu Arg Asn Cys Val Val Cys Gly Asp
 33 1 5 10 15
 35 Gln Ala Thr Gly Tyr His Phe Asn Ala Leu Thr Cys Glu Gly Cys Lys
 36 20 25 30
 38 Gly Phe Phe Arg Arg Thr Val Ser Lys Ser Ile Gly Pro Thr Cys Pro
 39 35 40 45
 41 Phe Ala Gly Ser Cys Glu Val Ser Lys Thr Gln Arg Arg His Cys Pro
 42 50 55 60
 44 Ala Cys Arg Leu Gln Lys Cys Leu Asp Ala Gly Met Arg Lys Asp Met
 45 65 70 75 80
 47 Ile Leu Ser Ala Glu Ala Leu Ala Leu Arg Arg Ala Lys Gln Ala Gln
 48 85 90 95
 50 Arg Arg Ala Gln Gln Thr Pro Val Gln Leu Ser Lys Glu Gln Glu Glu
 51 100 105 110
 53 Leu Ile Arg Thr Leu Leu Gly Ala His Thr Arg His Met Gly Thr Met
 54 115 120 125
 56 Phe Glu Gln Phe Val Gln Phe Arg Pro Pro Ala His Leu Phe Ile His
 57 130 135 140
 59 His Gln Pro Leu Pro Thr Leu Ala Pro Val Leu Pro Leu Val Thr His
 60 145 150 155 160
 62 Phe Ala Asp Ile Asn Thr Phe Met Val Leu Gln Val Ile Lys Phe Thr
 63 165 170 175
 65 Lys Asp Leu Pro Val Phe Arg Ser Leu Pro Ile Glu Asp Gln Ile Ser
 66 180 185 190
 68 Leu Leu Lys Gly Ala Ala Val Glu Ile Cys His Ile Val Leu Asn Thr
 69 195 200 205
 71 Thr Phe Cys Leu Gln Thr Gln Asn Phe Leu Cys Gly Pro Leu Arg Tyr

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/760,364

DATE: 08/17/2001

TIME: 12:09:45

Input Set : A:\-41-1.app

Output Set: N:\CRF3\08162001\I760364.raw

```

72      210      215      220
74 Thr Ile Glu Asp Gly Ala Arg Val Gly Phe Gln Val Glu Phe Leu Glu
75 225      230      235      240
77 Leu Leu Phe His Phe His Gly Thr Leu Arg Lys Leu Gln Leu Gln Glu
78      245      250      255
80 Pro Glu Tyr Val Leu Leu Ala Ala Met Ala Leu Phe Ser Pro Asp Arg
81      260      265      270
83 Pro Gly Val Thr Gln Arg Asp Glu Ile Asp Gln Leu Gln Glu Glu Met
84      275      280      285
86 Ala Leu Thr Leu Gln Ser Tyr Ile Lys Gly Gln Gln Arg Arg Pro Arg
87      290      295      300
89 Asp Arg Phe Leu Tyr Ala Lys Leu Leu Gly Leu Leu Ala Glu Leu Arg
90 305      310      315      320
92 Ser Ile Asn Glu Ala Tyr Gly Tyr Gln Ile Gln His Ile Gln Gly Leu
93      325      330      335
95 Ser Ala Met Met Pro Leu Leu Gln Glu Ile Cys Ser
96      340      345
99 <210> SEQ ID NO: 2
100 <211> LENGTH: 358
101 <212> TYPE: PRT
102 <213> ORGANISM: Mus musculus
104 <220> FEATURE:
105 <223> OTHER INFORMATION: mouse constitutive androstane receptor (CAR) beta 1
106      (mCARbeta1, mCAR1)
108 <400> SEQUENCE: 2
109 Met Thr Ala Met Leu Thr Leu Glu Thr Met Ala Ser Glu Glu Glu Tyr
110 1      5      10      15
112 Gly Pro Arg Asn Cys Val Val Cys Gly Asp Arg Ala Thr Gly Tyr His
113      20      25      30
115 Phe His Ala Leu Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Thr
116      35      40      45
118 Val Ser Lys Thr Ile Gly Pro Ile Cys Pro Phe Ala Gly Arg Cys Glu
119      50      55      60
121 Val Ser Lys Ala Gln Arg Arg His Cys Pro Ala Cys Arg Leu Gln Lys
122 65      70      75      80
124 Cys Leu Asn Val Gly Met Arg Lys Asp Met Ile Leu Ser Ala Glu Ala
125      85      90      95
127 Leu Ala Leu Arg Arg Ala Arg Gln Ala Gln Arg Arg Ala Glu Lys Ala
128      100      105      110
130 Ser Leu Gln Leu Asn Gln Gln Gln Lys Glu Leu Val Gln Ile Leu Leu
131      115      120      125
133 Gly Ala His Thr Arg His Val Gly Pro Leu Phe Asp Gln Phe Val Gln
134      130      135      140
136 Phe Lys Pro Pro Ala Tyr Leu Phe Met His His Arg Pro Phe Gln Pro
137 145      150      155      160
139 Arg Gly Pro Val Leu Pro Leu Leu Thr His Phe Ala Asp Ile Asn Thr
140      165      170      175
142 Phe Met Val Gln Gln Ile Ile Lys Phe Thr Lys Asp Leu Pro Leu Phe
143      180      185      190

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/760,364

DATE: 08/17/2001

TIME: 12:09:45

Input Set : A:\-41-1.app

Output Set: N:\CRF3\08162001\I760364.raw

```

145 Arg Ser Leu Thr Met Glu Asp Gln Ile Ser Leu Leu Lys Gly Ala Ala
146      195      200      205
148 Val Glu Ile Leu His Ile Ser Leu Asn Thr Thr Phe Cys Leu Gln Thr
149      210      215      220
151 Glu Asn Phe Phe Cys Gly Pro Leu Cys Tyr Lys Met Glu Asp Ala Val
152 225      230      235      240
154 His Ala Gly Phe Gln Tyr Glu Phe Leu Glu Ser Ile Leu His Phe His
155      245      250      255
157 Lys Asn Leu Lys Gly Leu His Leu Gln Glu Pro Glu Tyr Val Leu Met
158      260      265      270
160 Ala Ala Thr Ala Leu Phe Ser Pro Asp Arg Pro Gly Val Thr Gln Arg
161      275      280      285
163 Glu Glu Ile Asp Gln Leu Gln Glu Glu Met Ala Leu Ile Leu Asn Asn
164      290      295      300
166 His Ile Met Glu Gln Gln Ser Arg Leu Gln Ser Arg Phe Leu Tyr Ala
167 305      310      315      320
169 Lys Leu Met Gly Leu Ala Asp Leu Arg Ser Ile Asn Asn Ala Tyr
170      325      330      335
172 Ser Tyr Glu Leu Gln Arg Leu Glu Glu Leu Ser Ala Met Thr Pro Leu
173      340      345      350
175 Leu Gly Glu Ile Cys Ser
176      355
179 <210> SEQ ID NO: 3
180 <211> LENGTH: 286
181 <212> TYPE: PRT
182 <213> ORGANISM: Mus musculus
184 <220> FEATURE:
185 <223> OTHER INFORMATION: mouse constitutive androstane receptor (CAR) beta 2
186      (mCARbeta2, mCAR2)
188 <400> SEQUENCE: 3
189 Met Thr Ala Met Leu Thr Leu Glu Thr Met Ala Ser Glu Glu Glu Tyr
190 1      5      10      15
192 Gly Pro Arg Asn Cys Val Val Cys Gly Asp Arg Ala Thr Gly Tyr His
193      20      25      30
195 Phe His Ala Leu Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Thr
196      35      40      45
198 Val Ser Lys Thr Ile Gly Pro Ile Cys Pro Phe Ala Gly Arg Cys Glu
199      50      55      60
201 Val Ser Lys Ala Gln Arg Arg His Cys Pro Ala Cys Arg Leu Gln Lys
202 65      70      75      80
204 Cys Leu Asn Val Gly Met Arg Lys Asp Met Ile Leu Ser Ala Glu Ala
205      85      90      95
207 Leu Ala Leu Arg Arg Ala Arg Gln Ala Gln Arg Arg Ala Glu Lys Ala
208      100      105      110
210 Ser Leu Gln Leu Asn Gln Gln Gln Lys Glu Leu Val Gln Ile Leu Leu
211      115      120      125
213 Gly Ala His Thr Arg His Val Gly Pro Leu Phe Asp Gln Phe Val Gln
214      130      135      140
216 Phe Lys Pro Pro Ala Tyr Leu Phe Met His His Arg Pro Phe Gln Pro

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/760,364

DATE: 08/17/2001

TIME: 12:09:45

Input Set : A:\-41-1.app

Output Set: N:\CRF3\08162001\I760364.raw

```

217 145          150          155          160
219 Arg Gly Pro Val Leu Pro Leu Leu Thr His Phe Ala Asp Ile Asn Thr
220          165          170          175
222 Phe Met Val Gln Gln Ile Ile Lys Phe Thr Lys Asp Leu Pro Leu Phe
223          180          185          190
225 Arg Ser Leu Thr Met Glu Asp Gln Ile Ser Leu Leu Lys Gly Ala Ala
226          195          200          205
228 Val Glu Ile Leu His Ile Ser Leu Asn Thr Thr Phe Cys Leu Gln Thr
229          210          215          220
231 Glu Asn Phe Phe Cys Gly Pro Leu Cys Tyr Lys Met Glu Asp Ala Val
232 225          230          235          240
234 His Ala Gly Phe Gln Tyr Glu Phe Leu Glu Ser Ile Leu His Phe His
235          245          250          255
237 Lys Asn Leu Lys Gly Leu His Leu Gln Glu Pro Glu Tyr Val Leu Met
238          260          265          270
240 Ala Ala Thr Ala Leu Phe Ser Pro Gly Phe Cys Met Gln Ser
241          275          280          285

```

244 <210> SEQ ID NO: 4

245 <211> LENGTH: 492

246 <212> TYPE: DNA

247 <213> ORGANISM: Mus musculus

249 <220> FEATURE:

250 <223> OTHER INFORMATION: murine CARbeta genomic sequence - Section A,
251 portion of CARbeta intron sequence in left arm of

252 targeting construct

254 <220> FEATURE:

255 <221> NAME/KEY: modified_base

256 <222> LOCATION: (74)↓

257 <223> OTHER INFORMATION: n = g, a, c or t

259 <400> SEQUENCE: 4

```

260 aaaattttacc caacatagat ttatctaata taattcctat ctgcagaaca tccaaataact 60
W--> 261 ttggaaatta ttttttgggtg ttgtagctgt ttgaatgtaa acatatattc aaaaaaactc 120
262 ttcattggtga tgtagcattg ggcaagctat gaggatacct acttctggtt atttactaaa 180
263 agttgatagc caggcagtggt tggcacacac ctttaatccc agcacttggg aggcagaggc 240
264 aggtggaatt atgagtttga ggccagcctg gtctacagag tgggttcaag gtcagccagg 300
265 gctacacaga gaaacctgt ctcaaaaaga aggaggagga ggaggaaaga ggaagaggag 360
266 gaagaagatc ttttgttttg agatagcata cagtgtaaaat ttcggtttct ttagcaactc 420
267 agttgtgtca catgatgtct ttctggaagc tgtcttgtga gcagacatgt gatgtttatc 480
268 acaatagaaa gc 492

```

271 <210> SEQ ID NO: 5

272 <211> LENGTH: 1779

273 <212> TYPE: DNA

274 <213> ORGANISM: Mus musculus

276 <220> FEATURE:

277 <223> OTHER INFORMATION: murine CARbeta genomic sequence - Section B,
278 portion of CARbeta genomic sequence 5' to Section A

280 <400> SEQUENCE: 5

```

281 aaagagggtca tcaggcttgg cagcaagtgc ctttgcttac cgagtcttta caccagctcc 60
282 accgtggttt ttgagacagt ctccactgg actggatttc agcaagaaag ctaggcttgc 120

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/760,364

DATE: 08/17/2001

TIME: 12:09:45

Input Set : A:\-41-1.app

Output Set: N:\CRF3\08162001\I760364.raw

```

283 cttcttgtct ctgcctcctt ggcattggaa ttatgagttg ttccaccgtg ccatttttaa 180
284 aaatgtaggt tctaggaatt aaactcggtt ctcggtgctt atatagttag tactttacag 240
285 agggagtcac cttgccagca cctagaattc acttttatcc atatcccagt ctcctccagt 300
286 aagaaagtgg gateccttct agtggttacac ctaagttctt agttggatac cgaagtcttt 360
287 tttttaacag atctctgggg ctcagaaggc aagagctcct tgcagaggat ttaacctcaa 420
288 ttcttagtac tcaacttgcc agctcataac tgcctataac tctagtccca gaagatcaga 480
289 cattgtcctc tgatctctgt ggtactagg tatatacatt taaaaaaaat caataaaaaa 540
290 tttaaaaaaa gaaaagaaaa agaaagaaag aaaatccttt gggagcctgg tataattgtt 600
291 atagctacct tttttttttt tttttttttt ttttttacca ttgcaaaact gcacgtgaaa 660
292 aagcttgcca tctctcccat tgttctctgg cttattcagg atccatgcaa aaaggggagt 720
293 gtagatttag cctaaagctc acccacaggg aaatcctcca ggagtctagt aagcagcagc 780
294 ttttaatgag tcatgaggtc ctggccctc cccatctgcc accaaccaac acttctcggg 840
295 catgctagga acccccaccc caccacacac ccacaccag gtctttgccc tgggtccaga 900
296 gtctgggtcc tacctacata tggcaccgag gatacctaga ggcccatgc aagagaaggc 960
297 ccttgttttc caggcaactaa ggaccgcagt ccctaattcc tggcagttcc tgagatctca 1020
298 aggaaagcag ggtcagcgag gaggcctggg gagaggaggc atcctacacc cgatcttctg 1080
299 gctgctgcc taagggaac aggtaggtaa tccgttggag gccagagaca aaaagcaaca 1140
300 tttttgcttt taatgtcctc agtgctgggg agcccggtgt caggctgggc agtcttggga 1200
301 agagattctg tagaggagag agaagagagt cctatggccc agtgctgatt ctcaactcct 1260
302 cccacattca ggagaccatg acagctatgc taacactaga aacctggcc agtgaagaag 1320
303 aatatgggcc gaggaactgt gtggtgtgtg gagaccgggc cacaggctat catttccacg 1380
304 cctgacttg tgagggtctc aagggtctct tcagggtgaat gcttctctcc caacagaaac 1440
305 aaccccgaca tttctatcag tccaccttta aacactggta cactctcaag ttataatcct 1500
306 cttgcagcta agctgcactg cccagtgtct agcactctca atcttctga ccacaacgca 1560
307 gtgtgaaact ggtgacctaa tgacaaggca ggtaaacat ttgtcccaga gacagagcct 1620
308 aagagtcaag aacacttgtg tagcacacac tacctgcaaa gcaccgagat gattgccaca 1680
309 cgagggttcc tgagtaacct tgtgttctca tgaaaacgct ccaactacct ctgaagacct 1740
310 ttgagcacag ctcagatgag tctgttgta aatcgatcc 1779
313 <210> SEQ ID NO: 6
314 <211> LENGTH: 485
315 <212> TYPE: DNA
316 <213> ORGANISM: Mus musculus
318 <220> FEATURE:
319 <223> OTHER INFORMATION: murine CARbeta genomic sequence - Section C,
320 CARbeta intron sequence in right arm of targeting
321 construct
323 <400> SEQUENCE: 6
324 tgcattgctt tctactgaag tgtatcacag atgaatatga gatcgacaga aagtgtgcag 60
325 ggateccctt gccatctgga aacacttaat tcaatgaagt cccaaggaag cctcagaaac 120
326 tctttcttcc ttctccttcc cttatctggg gaggtggagt ggccccaact gaagggatgg 180
327 ctgaaaagggtg ctgctgctgt ttctcaacag ctttgtcatc tctcttgcct gacacagtga 240
328 tactgtcagc agaagccctg gcattgcggc gagccagaca ggcacagcgg cgggcagaga 300
329 aagcatcttt gcaactgaat cagcagcaga aagaactggt ccagatcctc ctcggggcc 360
330 aactcgcca tgtgggcccc atgtttgacc agtttgtgca gttcaagggt agaacttaac 420
331 caggatgtga cctgggtacc tgaggaggta acccacagaa gaaggctatg ccctgatgga 480
332 ggaca 485
335 <210> SEQ ID NO: 7
336 <211> LENGTH: 8
337 <212> TYPE: PRT

```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/760,364

DATE: 08/17/2001

TIME: 12:09:46

Input Set : A:\-41-1.app

Output Set: N:\CRF3\08162001\I760364.raw

L:261 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4

L:386 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8